Evolutionary Systematics. 4 2020, 21–33 | DOI 10.3897/evolsyst.4.49537

PENSOFT

EVOLUTIONARY SYSTEMATICS

The phylogeny of *Empis* and *Rhamphomyia* (Diptera, Empididae) investigated using UCEs including an over 150 years old museum specimen

Caroline Rhodén[1], Emma Wahlberg[1,2]

1 *Department of Zoology, Stockholm University, Stockholm, Sweden*

2 *Department of Zoology, Swedish Museum of Natural History, Stockholm, Sweden*

http://zoobank.org/58DF97B8-3FE5-4341-86E6-4BD5DDE777E0

Corresponding author: *Emma Wahlberg* (emma.wahlberg@nrm.se)

Academic editor: *M. Husemann* ♦ **Received** 19 December 2019 ♦ **Accepted** 21 January 2020 ♦ **Published** 4 February 2020

Abstract

The genera *Empis* Linneus, 1758 and *Rhamphomyia* Meigen, 1822 (Empidoidea, Empididae Latreille, 1809) are two large genera of flies commonly named dagger flies. They are widely distributed in the world with most species described from the Palearctic Region. *Empis* comprises about 810 described species and *Rhamphomyia* comprises about 610 described species, together they represent one third of the known species diversity in Empididae. Two recent studies on the phylogeny of the two genera using Sanger sequencing on a few genetic markers, did not support monophyly of them. In this study high throughput sequencing of target enriched molecular data of ultraconserved elements or UCEs was used to investigate the phylogenetic relationships of included representatives of the genera. This method has proven useful on old and dry museum specimens with high amounts of degraded DNA, which was also tested herein. For this purpose, a commercially synthesized bait kit has previously been developed for Diptera which this study was the first one to test. Three out of nine old and dry museum specimens were successfully sequenced, one with an age of at least 154 years. Higher DNA concentration yielded a greater number of reads. Analyses conducted in the study confirmed that both *Empis* and *Rhamphomyia* are non-monophyletic.

Key Words

Entomology, high throughput sequencing, next generation sequencing, systematics, taxonomy, target enrichment, UCE

Introduction

The family Empididae Latreille, in the superfamily Empidoidea Latreille, commonly known as dagger flies, is a family within Diptera consisting of around 3 051 known species in the world (Roskov et al. 2019). Dagger flies gets their vernacular name from the long and dagger-like piercing mouthparts. An older name is dance flies; however, this name is today assigned to the family Hybotidae Fallén, in the same superfamily. In both Empididae and Hybotidae many species form swarms where a typical mating ritual, which is perceived as a dance, takes place. Members in the subfamily Empidinae Latreille, constitute a high interspecific variation in mating rituals. (Cumming 1994; LeBas et al. 2003; Murray et al. 2018). The empidid tribe Empidini Latreille is highly diverse and consists of 14 genera spread all over the world, with a particular diversity in the Neotropical Region (Wiegmann et al. 2011). The two most species-rich genera in the tribe are the sister groups of *Empis* Linneus, 1758 and *Rhamphomyia* Meigen, 1822, and the majority of species of these two genera are described from the Palearctic Region (Watts et al. 2015). *Empis* constitute about 810 described species and *Rhamphomyia* about 610 described species according to 2019 Annual Checklist (CoL). Together they represent more than one third of all known Empididae species (Roskov et al. 2019). In an attempt to obtain a better overview of the diversity of the two genera

Copyright *Caroline Rhodén, Emma Wahlberg.* This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

several subgenera have been established. However, there is no clear number on how many subgenera there are but approximately 24 are given for *Empis* and 18 for *Rhamphomyia* (Chvála 1994; Poole and Gentili 1996; Yang et al. 2007; Saigusa 2012; Evenhuis and Pape 2019).

Previous studies have indicated that *Empis* and *Rhamphomyia* are non-monophyletic (Watts et al. 2016; Wahlberg and Johanson 2018). The study by Watts et al. (2015) found the genera to be polyphyletic, and that there is a Neotropical linage of *Empis* more closely related to the tribe Hilarini Collin, 1961, sister to Empidini. It was also hypothesized that further studies with additional sampling from the Palearctic and Nearctic regions and a larger molecular data set is necessary to resolve the phylogenetic relationships. Wahlberg and Johanson (2018) found *Empis* to be monophyletic except by two species of *Rhamphomyia* nested within it, implying non-monophyly of *Rhamphomyia*, but far from all subgenera were represented in the study and the specimens were mainly representatives of the Palearctic Region. The genera *Empis* and *Rhamphomyia* have several morphological resemblances involving a small head with large eyes, elongated mouth parts, long legs and an elongated abdomen with a high interspecific variability in male genitalia. The morphology of females has been much less studied compared to that of the males, and identification keys generally rely on male characters (Chvála 1994). Bothgenera can be morphologically distinguished from each other by the possession of a forked vein R_{4+5} in the wings in *Empis*, a feature lacking in *Rhamphomyia*. In addition, *Empis* species have much longer mouthparts compared to *Rhamphomyia* (Chvála 1994; Watts et al. 2016).

Former studies on the phylogeny of *Empis* and *Rhamphomyia* focused on either morphological or molecular data from traditional Sanger sequencing on only few genetic markers (Chvála 1994; Watts et al. 2016; Wahlberg and Johanson 2018. In this study we use a high throughput sequencing (HTS) method named target enrichment. This method allows for hundreds of genetic markers to be analysed at a less time and money expense per marker and specimen compared to the traditional Sanger sequencing. The targeted genetic markers focused on herein are called UCEs, ultraconserved elements. Which has proven useful for resolving phylogenies of less distant taxa of insects (Faircloth et al. 2012; Faircloth et al. 2014; Blaimer et al. 2016; Ješovnic et al. 2017; Van Dam et al. 2017). Another advantage of using this method is that it is useful on old and dried specimens which otherwise can pose a problem when using traditional Sanger sequencing due to the fragmentation of old DNA (Blaimer et al. 2016). For this purpose, commercially synthesized bait set which is complementary to the targeted UCEs sequences can be used to collect UCE data and its highly variable flanking regions adjacent to the UCE loci. (Faircloth et al. 2012). Until this point UCEs have not yet been tested on dipterans, this study is the first to evaluate the application of UCEs on Diptera (Ultraconserved 2017). Amplification of the cytochrome oxidase subunit I (COI) barcode gene is also performed in order to evaluate if the extractions went well and to validate species determinations in The Barcode of Life Database (BOLD) (Ratnasingham and Hebert 2007). In this study we include both relatively new material kept in ethanol sampled between 1995 and 2015 and older, dry museum specimens sampled between the years 1843 and 1993. The molecular data from the high throughput sequencing is used to investigate the phylogenetic relationship of the genera *Empis* and *Rhamphomyia*.

Methods

Material

A total of 48 taxa, 23 of *Empis* and 21 of *Rhamphomyia*, representing twelve subgenera of *Empis* and nine of *Rhamphomyia* were sampled from the collections at the Swedish Museum of Natural History in Stockholm (NHRS). The subgenera of *Empis* sampled are *Xanthempis* Bezzi, 1909, *Kritempis* Collin, 1926, *Empis s. str.*, *Euempis* Frey, 1953, *Anacrostichus* Bezzi, 1909, *Platyptera* Meigen, 1803, *Coptophlebia*, *Lissempis* Bezzi, 1909, *Polyblepharis* Bezzi, 1909, *Leptempis* Collin, 1926, *Pachymeria* Stephens, 1829, *Planempis* Frey,1953; and for *Rhamphomyia* are; *Aclonempis*, *Amydroneura* Collin, 1926, *Collinaria* Frey, 1950, *Eorhamphomyia* Frey, 1950, *Holoclera* Schiner, 1860, *Lundstroemiella* Frey, 1922, *Megacyttarus* Bigot, 1880, *Pararhamphomyia* Frey, 1922 and *Rhamphomyia* Meigen, 1822. Most of the samples were kept in 80% ethanol and were collected in Sweden through the Swedish Malaise Trap Project (SMTP) (Karlsson et al. 2005) between 2003 and 2006. Nine pinned specimens from 1843 to 1993 were included from the dry collection of the NHRS. Unidentified specimens were determined using the keys in Chvála (1994) and Collin (1961). Determined species were validated using COI barcodes where reference data was available in The Barcode of Life Data Systems (BOLD) (Ratnasingham and Hebert 2007). Four specimens from three genera from two tribes within Empididae were chosen as outgroup; Hilarini (*Hilara cornicula* Loew, 1873, *H. flavipes* Meigen, 1822) and Chelipodini Hendel, 1936 (*Chelipoda* sp. Macquart, 1823 and *Phyllodromia melanocephala* (Fabricius, 1794). The voucher numbers, collection data, and author for each species, are listed in Table 1.

DNA extraction and COI barcode amplification

For DNA extraction the KingFisher™ Cell and Tissue DNA Kit (Thermo Scientific, USA) was used together with KingFisher™ Duo (Thermo Scientific, USA) extraction robot following the manufacturer's protocols. For extractions of large specimens one leg was removed from the body, for medium sized specimens the abdomen was removed and for small ones the whole animal was used. Lysis was performed in 56 °C overnight. After extraction the body part was returned to the specimen. For pinned

material only oneg was used, due to restrictions from the museum. All extracted material is kept in 80% ethanol as vouchers at the Swedish Museum of Natural History (NHRS). For pinned material a voucher code was attached to the pin and the specimen was returned to the collection. Amplification of the COI barcode gene was performed for all samples in order to evaluate if the extractions went well and to validate species determinations. The PCR reactions were carried out with a 25 µl reaction containing Ready-To-Go PCR Beads (Amersham Biosciences, Great Britain) and 1 µl of each primer, 2 µl DNA template and 21 µl ddH_20 for each sample. The primers used were LCO1490 and HCO2198 (Folmer et al. 1994). The amplification program included 95 °C for 5 min, followed by 40 cycles of 95 °C for 30 sec, 50 °C for 30 sec and 72 °C for 50 sec and a final step at 72 °C for 8 min. To determine whether the amplification was successful the PCR product was inspected using gel electrophoresis. The successful PCR products were purified using Exo-Fast (Qiagen, Germany), and then sent to Macrogen Inc (Netherlands) for Sanger sequencing. Barcode sequences were deposited at NCBI GenBank, accession numbers are given in Table 1.

Library construction and target enrichment

The DNA concentration and fragmentation in each extraction of the nine old and dry samples was measured using Qubit (Thermo Fischer Scientific, USA) and BioAnalyzer (Agilent, USA). The samples from newer specimens were measured and compared to the older samples. The newer samples were fragmented on a Covaris sonicator, at SciLife Lab (Solna), to the target fragment length of 500–600 bp. Libraries for each sample were constructed following a modified version of the Meyer and Kircher (2010) protocol for Illumina sequencing, using magnetic AMPure beads for cleaning steps. The modified protocol does not contain the step of fragmentation and purification of sample DNA and the temperature profile of the PCR reactions is slightly different. This modification has been developed in house to fit museum samples. The protocols used for library preparation and amplification following hybridization are available in Meyer and Kircher (2010). Adapters used were IS1_adapter_P5.F, IS2_adapter_P7.F and IS3_adapter_P5+P7.R. The libraries were amplified with dual index primers. Before hybridization step the DNA concentration was measured again and fragment size distribution inspected on BioAnalyser. Size selection and purification of libraries was carried out using AmPure XP (Agencourt, France) beads, with a 1.8X ratio, and thereafter pooled in equimolar amounts into 8 pools for the following hybridization step. The hybridization was conducted following the myBaits Hybridization Capture for Targeted NGS version 4.01 protocol and the myBaits UCE Diptera 2.7Kv1 baits kit constructed by Faircloth (2017) was used (synthesized by Arbor Biosciences, USA). The hybridization was conducted in 65 °C for 18 hours. KAPA HiFi HotStart was used for library amplification with the primers IS5_reamp.P5 forward library primer (10 µM) and IS6_reamp.P7 reverse library primer (10 µM). After hybridization the 8 pools were pooled into one pool in equimolar amounts and sequenced on an Illumina MiSeq v3 2x300bp pair-end platform at SciLife Lab (Solna, Sweden). Raw reads were deposited at NCBI Sequence Read Archive (SRA) as a BioProject, accession number PRJNA596621.

Assembly and alignment of UCE sequences

Demultiplexed reads were quality checked and filtered using the pre-processing tool fastp (Chen et al. 2018) with standard settings and base correction for paired end data. Using the base correction for paired end data also merged forward and reverse reads in one step. Assembly of paired reads were conducted using METASPADES (Nurk et al. 2017). The extraction of UCE sequences, alignment, cleaning and preparation of UCE data followed the PHYLUCE pipeline by Faircloth (2016). The extraction of UCE data was performed using the Diptera 2.7Kv1 probes (Faircloth 2017). Nucleotide based alignment was carried out in MAFFT v7 (Katoh and Standley 2013) with no trimming. Edge and internal trimming of the alignments was conducted outside the pipeline with TRIMAL v.1.2 (Capella-Gutierrez et al. 2009) to remove poorly aligned or ambiguous sites. The alignments were optimized prior to the phylogenetic analysis by finding the best partitioning scheme and substitution models. To create a table of partitions for UCE data PFINDERUCE-SWSC-EN v1.0.0 (Tagliacollo and Lanfear 2018) was used to identify the conservative core and variable flanking regions. Partition scheme and substitution model test was performed in PARTITIONFINDER v2.1.1 (Lanfear et al. 2012) with the options for rcluster and RAxML algorithms. A final dataset of 70% completeness was created for further phylogenetic analysis.

Phylogenetic analysis

Bayesian inference was performed on the partitioned dataset using MRBAYES v3.2.6 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003) and set with the substitution models generated by PartitionFinder for each partition. The following settings were used for the analysis; 4 chains, 2 runs, 100 million generations, sampling frequency of 10 000, the temperature was set to 0.11, burnin of 25%. The log files were inspected in TRACER v1.7.1 (Rambaut et al. 2018) to determine the burnin and the effective sample size (ESS). Maximum likelihood analysis was performed in IQTREE v1.6.10 (Nguyen et al. 2015) on the CIPRES Science Gateway v.3.3 (Miller et al. 2010) with a non-parametric bootstrap analysis with 500 replicates. The resulting trees were viewed and edited in FIGTREE v1.4.4 (Rambaut 2014) and ADOBE ILLUSTRATOR v24.0.1. The trees were rooted on the two species of Hilarini, *Hilara cornicula* (AG7C) and *Hilara flavipes* (AG8C).

Table 1. List of specimens including voucher number, collection data, geographic origin and accession number of COI barcode sequence in GenBank (if available). SMTP referring to Swedish Malaise Trap Project (Karlsson et al. 2005).

Taxon	Voucher	Locality and collector	Collection date	GenBank accession number (COI)
Chelipoda sp.	AH7C	NEW CALEDONIA: Province Sud, Platou de Dogny, source of Dogny River, about 500 m SE summit of Platou de Dogny, –21.613588; 165.883401. Leg. K. A. Johanson.	25 Nov.–16 Dec. 2003	MN868983
E. (*Coptophlebia*) *albinervis* Meigen, 1822	AA9C	SWEDEN: Öland, Mörbylånga kommun, Gamla Skogsby (Kalkstad), 56.616700; 16.507617. Leg. SMTP.	3–20 Jul. 2006	–
E. (*Empis*) *bicuspidata* Collin, 1927	AB1C	SWEDEN: Torne lappmark, Kiruna kommun, Abisko nationalpark, Nuolja, 68.359492; 18.719197. Leg. SMTP.	26 Jun.–15 Jul. 2006	MN868986
E. (*Platyptera*) *borealis* Linnaeus, 1758	AA8C	SWEDEN: Torne lappmark, Kiruna kommun, Abisko nationalpark, Nuolja, 68.359492; 18.719197. Leg. SMTP.	11–17 Aug. 2005	MN868968
E. (*Leptempis*) *grisea* Fallén, 1816 Old specimen	AE3C	SWEDEN: Skåne. Leg. C. H. Boheman.	1852 or 1865	–
E. (*Coptophlebia*) *hyalipennis* Fallén, 1816 Old specimen	AE5C	SWEDEN: Skåne. Leg. C. H. Boheman.	1852 or 1865	–
E. (*Planempis*) *latro* Frey, 1953	AF4C	JAPAN: Hyōgo Prefecture, Arimafuji Park, 34.9194; 135.2356. Leg. Japan 2011 Exp.	12 Apr. 2011	MN868961
E. (*Kritempis*) *livida* Linnaeus, 1758	AA3C	SWEDEN: Östergötland, Ödeshögs kommun, Omberg, bokskogsreservatet (beech forest), 58.297183; 14.634817. Leg. SMTP.	5–19 Jul. 2005	MN868964
E. (*Anacrostichus*) *lucida* Zetterstedt, 1838	AA4C	SWEDEN: Torne lappmark, Kiruna kommun, Abisko nationalpark, Nuolja, 68.359492; 18.719197. Leg. SMTP.	1–13 Jul. 2005	MN868972
E. (*Leptempis*) *nigricans* Meigen, 1804 Old specimen	AD9C	SWEDEN: Skåne. Leg. P. F. Wahlberg.	1846	–
E. (*Lissempis*) *nigritarsis* Meigen, 1804	AD1C	SWEDEN: Öland, Mörbylånga, Gamla Skogsby (Kalkstad), mixed deciduous forest. Leg. M. & C. Jaschhof.	1–25 May 2014	MN868979
E. (*Anacrostichus*) *nitida* Meigen, 1804 Old specimen	AE1C	SWEDEN: Jämtland, Undersåker. Leg. O. Ringdahl.	17 Jun. 1914	–
E. (*Empis*) *nuntia* Meigen, 1838	AB5C	SWEDEN: Öland, Mörbylånga, Lilla Vickleby Lunds NR, old oak forest. Leg. M. & C. Jaschhof.	1–27 May 2014	MN868989
E. (*Euempis*) *picipes* Meigen, 1804 Old specimen	AE2C	SWEDEN:Södermanland, Stormossen. Leg. A. Orbe.	19 Apr. 1991	–
E. (*Empis*) *planetica* Collin, 1927	AA7C	SWEDEN: Uppland, Knivsta kommun, Rickebasta alsumpskog, western part, 59.734350; 17.720417. Leg. SMTP.	18 Jun.–6 Jul. 2003	MN868966
E. (*Euempis*) *tessellata* Fabricius, 1794	AA6C	SWEDEN: Småland, Gränna kommun, Lönnemålen, 58.048917; 14.573033. Leg. SMTP.	15 Jun.–1 Jul. 2005	MN868980
E. (*Xanthempis*) *trigramma* Wiedemann, 1822	AB2C	SWEDEN: Öland, Mörbylånga, Gamla Skogsby (Kalkstad), mixed deciduous forest. Leg. M. & C. Jaschhof.	1–25 May 2014	MN868960
E. (*Xanthempis*) *univittata* Loew, 1867	AB3C	SWEDEN: Öland, Mörbylånga, Kalkstad NR, mixed deciduous forest. Leg. M. & C. Jaschhof.	27 May–27 Jun. 2014	MN868951
Empis sp. 1	AD5C	RUSSIA: Chukota Autonomous Okrug, Chaunsky, Ajon Island, 70 m from base camp, 69.5840; 168.6955. Leg. P. Mortensen.	11–24 Jul. 2015	MN868973
Empis sp. 2	AD7C	JAPAN: Ehime Prefecture, Matsuyama–shi, Ehime University Forest. Leg. Japan 2011 Exp.	2 May 2011	MN868957
Empis sp. 3	AF2C	GREECE: East Macedonia, Paranesti by Nestos River, east bank of river north of road to Drama. Leg. P. Lindskog & B. Viklund.	3–5 May 1995	MN868970
Empis sp. 5	AG9C	CHILE: Region de los Lagos, Isla Grande de Chiloé Rio Melilebú, along road between Tebuhueico and Hullinco, 5.4 km S crosspoint to Curaco, river, 100 m upstream bridge, –42.7181; –73.8965. Leg. K. A. Johanson.	6 Jan. 2006	MN868969
Empis sp. 6	AH2C	NEW CALEDONIA: Province Sud, Platou de Dogny, source of Dogny River, about 500 m SE summit of Platou de Dogny, –21.613588; 165.883401. Leg. K. A. Johanson.	25 Nov.–16 Dec. 2003	

Taxon	Voucher	Locality and collector	Collection date	GenBank accession number (COI)
Hilara cornicula	AG7C	SWEDEN: Uppland, Älvkarleby kommun, Älvkarleby kommun, Båtfors, between Milsten and Båtforstorpet, 60.46065; 17.317817. Leg. SMTP.	27 Jun.–1 Jul. 2004	MN868950
Hilara flavipes	AG8C	SWEDEN: Ångermanland, Örnsköldsviks kommun, Skuleskogen, Långrå, 63.088717; 18.498383. Leg. SMTP.	5–25 Jul. 2004	MN868984
Phyllodromia melanocephala	AH5C	SWEDEN: Hälsingland, Hudiksvalls kommun, Stensjön–Lomtjärn, Stensjön, Marsh pine wood close to bog, 62.140333; 16.286100. Leg. SMTP.	8–23 Apr. 2005	–
R. (*Megacyttarus*) *anomalina* Zetterstedt, 1838	AC5C	SWEDEN: Ly, Sorsele kommun, Ammarnäs, Vindelfjällens naturreservat, Tjulträstklaspen. 65.966783; 16.060500. Leg. SMTP.	28 Jun.–15 Jul. 2004	–
R. (*Megacyttarus*) *crassirostris* (Fallén, 1816)	AC2C	SWEDEN: Lycksele lappmark, Sorsele kommun, Ammarnäs, Vindelfjällens naturreservat, Tjulträstklaspen, 65.966783; 16.060500. Leg. SMTP.	28 Jun.–15 Jul. 2004	MN868975
R. (*Pararhamphomyia*) *curvula* Frey, 1913	AC9C	SWEDEN: Lycksele lappmark, Sorsele kommun, Ammarnäs, Vindelfjällens naturreservat, Tjulträstklaspen, 65.966783; 16.060500. Leg. SMTP.	28 Jun.–15 Jul. 2004	
R. (*Lundstroemiella*) *dudai* Oldenberg, 1927	AB7C	SWEDEN: Ångermanland, Örnsköldsviks kommun, Skuleskogen, Långrå, 63.088717; 18.498383. Leg. SMTP.	5–25 Aug. 2004	MN868958
R. (*Amydroneura*) *erythrophthalma* Meigen, 1830	AD3C	SWEDEN: Skåne, Ystad kommun, Sandhammaren, Järahusen, 55.403781; 14.199936. Leg. SMTP.	26 Sep. 2005–10 Feb. 2006	MN868965
R. (*Pararhamphomyia*) *fascipennis* Zetterstedt, 1838	AB9C	SWEDEN: Värmland, Munkfors kommun, Ransäter, Ransberg Herrgård, 59.790442; 13.415169. Leg. SMTP.	22 May–5 Jun. 2005	MN868954
R. (*Amydroneura*) *gibba* (Fallén, 1816) Old specimen	AF1C	SWEDEN: Skåne. Leg. C. H. Boheman.	1852 or 1865	–
R. (*Lundstroemiella*) *hybotina* Zetterstedt, 1838	AD2C	SWEDEN: Hälsingland, Hudiksvalls kommun, Stensjön–Lomtjärn, Stensjön, 62.140333; 16.286100. Leg. SMTP.	14–27 Jul. 2005	MN868949
R. (*Holoclera*) *nigripennis* (Fabricius, 1794)	AC7C	SWEDEN: Öl. Mörbylånga kommun, Västerstads almlunds naturreservat, old elm forest. 56.427307; 16.421942. Leg. SMTP.	15 May–9 Jul. 2014	MN868962
R. (*Rhamphomyia*) *nigrita* Zetterstedt, 1838	AD6C	RUSSIA: Chukota Autonomous Okrug, Chaunsky, Ajon Island, 70 m from base camp, 69.5840; 168.6955. Leg. P. Mortensen.	11–25 Jul. 2015	MN868988
R. (*Collinaria*) *nitidula* Zetterstedt, 1842 Old specimen	AE8C	SWEDEN: Torne lappmark, Kiruna kommun, Abisko nationalpark. Leg. O. Ringdahl.	30 Jun. 1918	–
R. (*Pararhamphomyia*) *pilifer* Meigen, 1838	AB8C	SWEDEN: Ångermanland, Örnsköldsviks kommun, Skuleskogen, Långrå, 63.088717; 18.498383. Leg. SMTP.	5–25 Aug. 2004	MN868990
R. (*Rhamphomyia*) *plumipes* (Meigen, 1804) Old Specimen	AE6C	SWEDEN: Lappland, Hemavan, Klippen. Leg. H. Bartsch.	4 Jul. 1993	–
R. (*Eorhamphomyia*) *spinipes* (Fallén, 1816) Old specimen	AE7C	SWEDEN: Lappland. Leg. N. J. Andersson.	1843	–
R. (*Rhamphomyia*) *sulcata* (Meigen, 1804)	AC8C	SWEDEN: Lycksele lappmark, Sorsele kommun, Ammarnäs, Vindelfjällens naturreservat, Tjulträstklaspen, 65.966783; 16.060500. Leg. SMTP.	28 Jun.–15 Jul. 2004	MN868982
R. (*Rhamphomyia*) *trilineata* Zetterstedt, 1859	AB6C	SWEDEN: Ångermanland, Örnsköldsviks kommun, Skuleskogen, Långrå, 63.088717; 18.498383. Leg. SMTP.	5–25 Aug. 2004	MN868987
R. (*Holoclera*) *umbripennis* Meigen, 1822	AC1C	SWEDEN: Torne lappmark, Kiruna kommun, Abisko nationalpark, Nuolja, 68.359492; 18.719197. Leg. SMTP.	1–13 Aug. 2005	MN868952
Rhamphomyia sp. 1	AD4C	RUSSIA: Chukota Autonomous Okrug, Chaunsky, Ajon Island, 70 m from base camp, 69.5840; 168.6955. Leg. P. Mortensen.	11–24 Jul. 2015	MN868956
Rhamphomyia sp. 2	AD8C	FRENCH GUIANA: Approuague–Kaw, Kaw Mountain, 4.5508; –52.1949. Leg. N. Jönsson.	4–12 Feb. 2007	MN868967
Rhamphomyia sp. 3	AF7C	JAPAN: Ehime Prefecture, Tōon–shi, Saragamine Range Prefectural Park, 33.71598; 132.8943. Leg. Japan 2011 Exp.	18 Apr. 2011	MN868963
Rhamphomyia sp. 4	AF8C	JAPAN: Ehime Prefecture, Tōon–shi, Saragamine Range Prefectural Park, 33.71598; 132.8943. Leg. Japan 2011 Exp.	18 Apr. 2011	MN868955

Results

Libraries, alignment and partitioning

Of the 48 specimens ten lack sufficient reads or target loci and were removed from the analyses. Five out of the nine old and dry samples were successfully aligned but only three of them, *E. picipes* (AE2C), *E. hyalipennis* (AE5C) and *R. gibba* (AF1C), with a satisfactory amount of data. In 35 of the 38 samples there was an increase in DNA concentration after library amplification. DNA concentration for the old and dry samples in the study range between 0.626 and 1.09 ng/µl before the library amplification. For newer material stored in ethanol the concentration range between 0.198 and 11.0 ng/µl. Following library amplification DNA concentration of the old samples range between 1.58 and 13.9 ng/µl and for the newer 0.612 and 46.8 ng/µl. The number of reads for each of the 38 specimens varies between 38 000 and 3 000 000. Seven out of the ten poorly sequenced samples have a lower number of reads than the rest, ranging from 24 to 28 000. However, three of them have reads ranging between 80 000 and 176 000. The measured DNA concentration before and after library construction is depicted for each taxon in Appendix 1 including the ten excluded samples. The dataset has 41 out of 48 specimens with enough loci represented. Six of the seven excluded specimens were old and dry. The number of aligned loci were 15, the alignment length was 7 394 bp and the number of informative sites were 1 900 bp. The number of partitions for the dataset was 21. Detailed partition schemes with models chosen for the partitions are summarized in Appendix 2.

Phylogenetic analysis

The phylogenetic analysis of the dataset generated trees with a total of 38 taxa; 36 ingroup taxa and two outgroup taxa (Figs 1, 2). The two specimens *E. grisea* (AE3C) and *E. nigricans* (AD9C) were old and dry museum samples and were removed from the analysis because their sequences were short and has long gaps and generated very long branches in the phylogenetic trees. The ESS-values range between 4800 and 7500 for separate runs, and 2700 to 7500 for combined runs. The two trees have some differences in topology; however, the Bayesian inference-tree has a higher support in general. In the Bayesian inference-tree (Fig. 1) 33 nodes have a posterior probability support above 93%. In the Maximum likelihood-tree (Fig. 2) 12 of the most recent nodes have a bootstrap support above 85. *Empis* and *Rhamphomyia* were divided into multiple well supported monophyletic groups scattered in the tree, leaving both non-monophyletic. The clades A, B and C marked in blue in Fig. 1 contain both genera. Clades A, B and C have high (> 93%) posterior probabilities in the Bayesian inference-tree, in the Maximum likelihood-tree Clade A has a support value of 100, Clade B has a support of 41 and Clade C is not present. In the Bayesian inference-tree Clade B includes five *Empis* species and three *Rhamphomyia* species with a support of 0.93.

Discussion

The application of target enrichment of UCEs on old and dry museum samples of Diptera was in general successful. Three of the old samples were sequenced *E.* (*Euempis*) *picipes* (AE2C), *R.* (*Rhamphomyia*) *plumipes* (AE6C) and *R.* (*Amydroneura*) *gibba* (AF1C) with sufficient data, but two specimens, *E.* (*Leptempis*) *nigricans* (AD9C) and *E.* (*Leptempis*) *grisea* (AE3C), had too little data coverage. When performing Sanger sequencing of COI none of the old and dry specimens were recovered. This was expected because old and dry samples usually have highly degraded DNA, and to conduct a successful PCR, preceding Sanger sequencing, it is necessary with sufficient good DNA quality (Lindahl 1993; Junqueira et al. 2002). A recently conducted study applying target enrichment using UCEs succeeded in analysis of a 121 year old museum specimen of carpenter bee (Blaimer et al. 2016). In this study *R.* (*Amydroneura*) *gibba* (AF1C) was sampled by Carl H. Boheman (1796–1868), either in 1852 or 1865, so this sample is at least 154 years old. The specimens of *E.* (*Euempis*) *picipes* (AE2C) and *R.* (*Rhamphomyia*) *plumipes* (AE6C) were sampled in 1991 and 1993 respectively. This result shows that there is a great possibility to utilize the DNA of the immense collections of old and dry specimens of Diptera at natural history museums in the world.

The average DNA concentration in this study (2.26–12.94 ng/µl) were within the same range as in Blaimer et al. (2016), with old samples (1.11–14.67 ng/µl,), whereas in Ješovnic et al. (2017), a study with newer samples, the DNA concentration were in a higher range (3.87–79 ng/ µl). Blaimer et al. (2016) found that DNA concentration and number of reads of old pinned specimens decreased with an increasing age of the sample. Most of the old samples in this study did not express a large increase in DNA concentration after library amplification and the number of reads were low compared to the majority of the newer samples. Comparing the number of reads of the 38 samples (38 000–3 000 000) to three other studies the samples range start from a much lower number. Samples in Blaimer et al. (2016) ranged between 70 256–3 479 137, samples in Ješovnic et al. (2017) ranged between 299 485–3 500 409, in Van Dam et al. (2017) samples ranged between 1 716 890–31 283 213. As stated above, these studies also had a generally higher DNA concentration than our specimens. The studies by Van Dam et al. (2017) and Ješovnic et al. (2017) did not have any samples as old as in the current study, which might explain the difference. But it could also be due to the amount of tissue sampled, from the old samples we were only allowed to extract one leg of quite small specimens. Possibly a larger amount of tissue would increase the DNA concentration and thereby the number of reads.

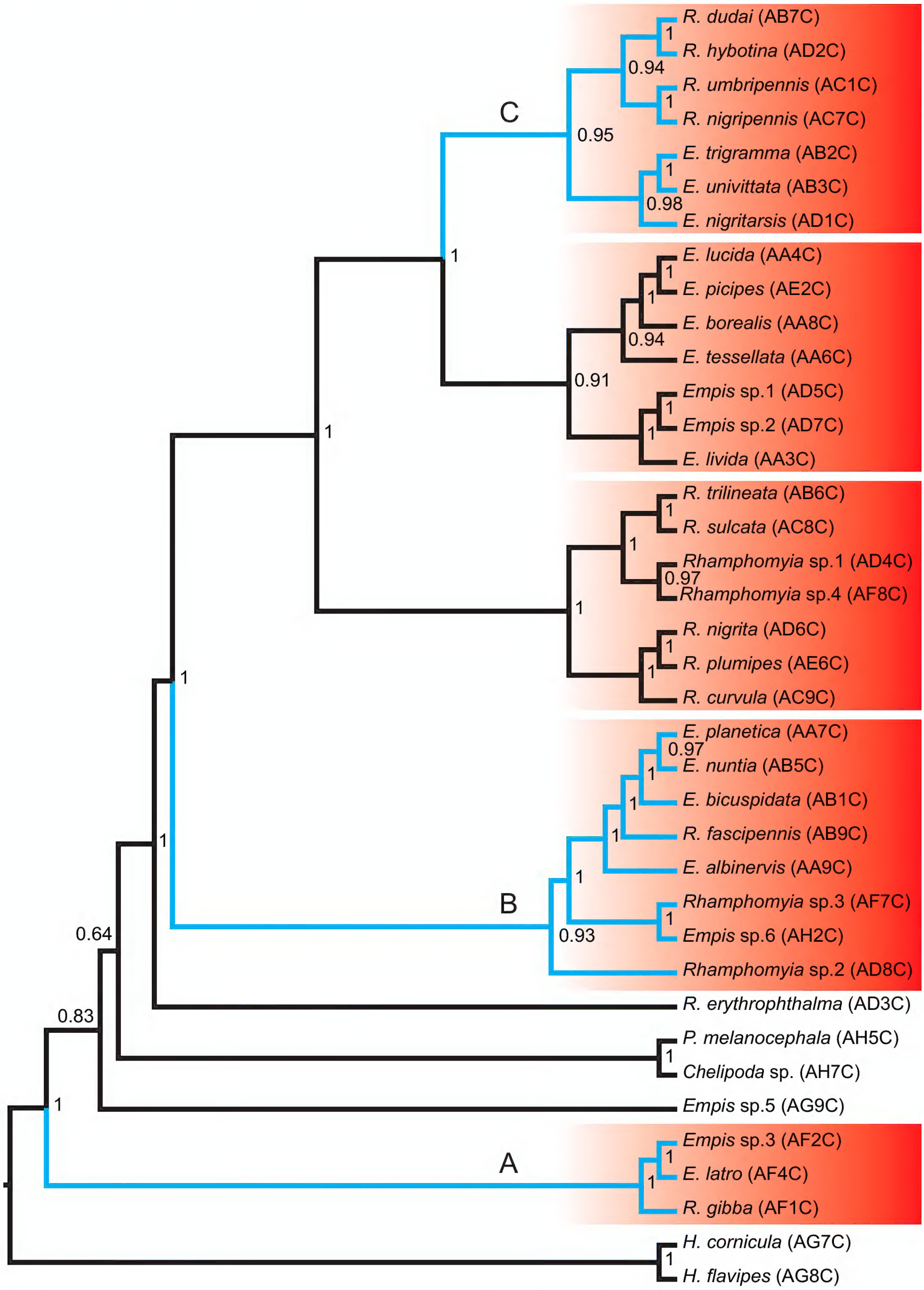

Figure 1. The majority rule tree of partitioned UCE data of 70% completeness, inferred with Bayesian inference in MRBAYES v3.2.6 with a burnin of 25%. Posterior probability values above 50% are depicted at nodes. Voucher numbers are provided in parenthesis for all taxa. The blue branches mark the Clade A, B and C which are clades of both genera forming monophyletic groups, the red boxes mark monophyletic clades.

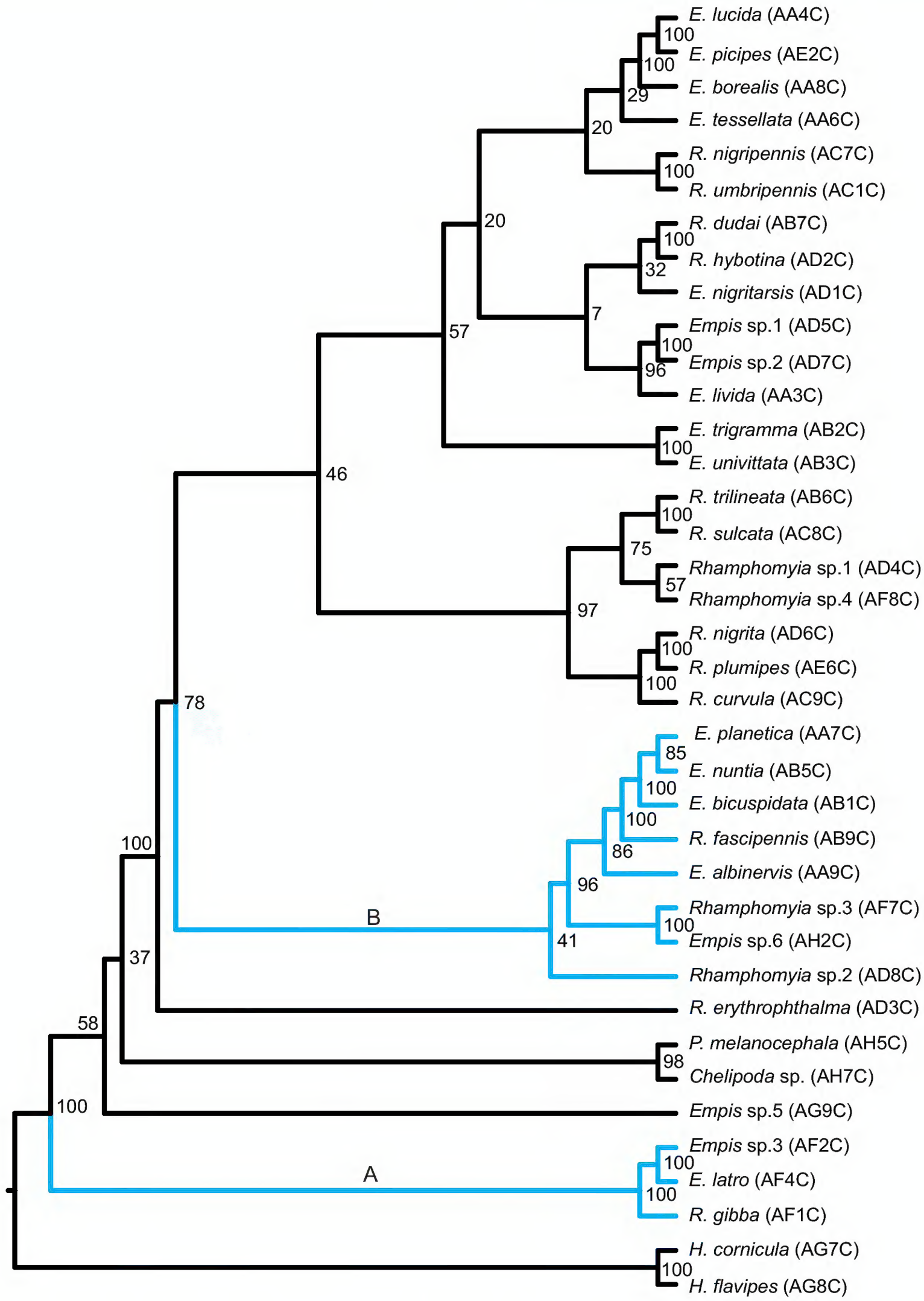

Figure 2. Maximum likelihood tree of partitioned UCE data of 70% completeness, inferred with IQTREE v1.6.10. Bootstrap values are depicted at nodes. Voucher numbers are provided in parenthesis for all taxa. The blue branches mark the Clade A and B which are clades of both genera forming monophyletic groups.

Other factors that might affect DNA concentrations and fragmentations are extraction protocols. This is the first study conducted on dipterans using target enrichment of UCEs. Further development of specially DNA extraction protocols might refine the methodology.

The phylogenetic analyses adopting Bayesian inference inferred a tree with high support values (Fig. 1). The genera were widely scattered in the tree which contradicts the hypothesis that the genera are monophyletic. What strengthens the non-monophyly even more is the three clades A, B and C depicted in the tree (Fig. 1). In clade B there is a subclade of *Rhamphomyia* sp. 3 (AF7C) and *Empis* sp. 6 (AH2C) with short branches, these two species belong to the *Rhamphomyia* subgenus *Aclonempis* and *Empis* subge-

Figure 3. *Empis* (*Empis*) *planetica* (voucher AA7C), with differing wing characters. The left wing possesses a R_{4+5} fork and the right wing lack the R_{4+5} fork.

nus *Coptophlebia*. These subgenera have been discussed by Chvála (1994), who suggested their monophyly if including the subgenus *Empis s. s*. All sampled individuals belonging to *Empis s. s.*, *Coptophlebia* and *Aclonempis* are in this study grouped in clade B. Chvála (1994) has also stated that the *Empis* subgenus *Lissempis* is more closely related to the *Rhamphomyia* subgenus *Lundstroemiella* than to any other *Empis* subgenus. In clade C the *Empis* subgenus *Lissempis* is more related to the *Empis* subgenus *Xanthempis* but is sister group to *Lundstroemiella*. The high posterior probability values in the Bayesian inference-tree confirms the non-monophyly of the genera previously suggested by Watts et al. (2016) based on analyses of Sanger sequenced data. In the study by Watts et al. (2016) geographic distribution was taken into account and it was found that there are two linages, one linage with Palearctic + Nearctic *Empis* and *Rhamphomyia* and one linage with Neotropical *Empis*. The two linages were recovered as sister groups and Neotropical *Empis* was more closely related to the Empidini genera *Lamprempis* Wheeler & Melander, 1901, *Opeatocerata* Melander, 1928, *Macrostomus* Wiedemann, 1817 and *Porphyrochroa* Melander, 1928 than to the other linage. Our sampling is mainly Palearctic; Sweden, Russia, Japan, Greece, but with additional taxa from French Guiana, Chile and New Caledonia. One of the two Neotropical species, the Chile-species *Empis* sp. (AG9C), is placed as a sister group to all other taxa except three Palearctic species *R.* (*Amydroneura*) *gibba* (AF1C), *Empis* sp. (AF2C) and *E.* (*Planempis*) *latro* (AF4C). This corresponds to the findings by Watts et al. (2016).

The species within the two genera are morphologically quite similar, and the traditional characters used to distinguish the genera are the wing venation and mouthpart length. However, there are exceptions. For example, species in the *Rhamphomyia* subgenera *Aclonempis* and *Vockerotempis*, Saigusa 2012 possess a long labrum much like those found in *Empis* species. Another important factor affecting stability of classification based on wing venation is that of intraspecific variation, even within the same exemplar, i.e. one wing having a R_{4+5} fork and the other lacking the fork (Chvála 1994). Such a case was found in this study, the species *E.* (*Empis*) *planetica* (AA7C) (Fig. 3). This species was placed as a sister taxon to *E.* (*Empis*) *nuntia* (AB5C) in clade B in the Bayesian inference and Maximum likelihood-tree. This raises the question of how reliable these morphological traits are for separating the genera and species. Lastly, another interesting finding is that the two Chelipodini species representing the outgroup together with *Hilara* are both nested within Empidinae as a sister group to all other species of the subfamily except the Chilean species and three Palearctic species *R.* (*Amydroneura*) *gibba* (AF1C), *Empis* sp. (AF2C) and *E.* (*Planempis*) *latro* (AF4C). Rooting on these two groups respectively did not change the tree topology. The trees were rooted on Hilarini, as according to the latest research of the superfamily Empidoidea which concluded that Chelipodini is more closely related to Empidini than Hilarini is (Wahlberg and Johanson 2018). The previous study by Watts et al. (2016) suggest that *Hilara* is a sister-group to *Empis* and *Rhamphomyia*. However, the support for the placement of Chelipodini in this study is low, 0.64.

Conclusion

The first-time application of the Diptera 2.7Kv1 probe kit (Faircloth 2017) for target enrichment using UCEs was successful regarding inferring phylogenies with high support in the Bayesian inference analysis. Sequences of five out of nine old museum samples were successfully aligned, however only three were good enough to be used in a phylogenetic analysis. For future studies we suggest increasing the tissue sampling on old material of Diptera to increase the chances of higher DNA concentration. However, this result

shows that it is possible to use the immense collections of old and dry Diptera samples for DNA studies. The application of this technique can reduce the sampling of new specimens which would be beneficial for the biodiversity. Exchange of specimens between natural history museums, universities and other collections can provide researchers with specimens from all over the world. The analyses performed well in this study and inferred *Empis* and *Rhamphomyia* as non-monophyletic. This corresponds with the studies of Watts et al. (2016) and Wahlberg and Johanson (2018) using Sanger sequencing (;).

Acknowledgements

Thanks to Rodrigo Esparza-Salas and Bodil Cronholm at the Molecular Systematics Laboratory (Swedish Museum of Natural History) for the valuable help with the laboratory work. Also, a big thank to SciLife Lab (Solna) for guidance and help during sample preparation. The major part of the specimens used in this study was provided by the Swedish Malaise Trap Project (SMTP). Additional specimens were collected by Niklas Apelqvist, Mathias and Catrin Jaschhof, Kjell Arne Johanson, Peter Mortensen, Pär Lindskog and Bert Viklund. Thanks to Yngve Brodin (curator) for the permission to use specimens from the Swedish Museum of Natural History collection. Thanks to Yngve Brodin and Mattias Forshage for valuable comments on the manuscript.

References

Bejerano G, Pheasant M, Makunin I, Stephen S, Kent WJ, Mattick JS, Haussler D (2004) Ultraconserved Elements in the Human Genome. Science 304(5675): 1321–1325. https://doi.org/10.1126/science.1098119

Blaimer BB, Lloyd MW, Guillory WX, Brady SG (2016) Sequence Capture and Phylogenetic Utility of Genomic Ultraconserved Elements Obtained from Pinned Insect Specimens. PLOS ONE 11(8): e0161531. https://doi.org/10.1371/journal.pone.0161531

Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25(15): 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ pre-processor. Bioinformatics 34(17): 884–890. https://doi.org/10.1093/bioinformatics/bty560

Chvála M (1994) The Empidoidea (Diptera) of Fennoscandia and Denmark. III Genus *Empis*. E.J Brill, Leiden, 192 pp.

Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2016) GenBank. Nucleic acids research 44(D1): D67–D72. https://doi.org/10.1093/nar/gkv1276

Collin JE (1961) British Flies: Empididae. Part II: Empididae. Cambridge University Press, Cambridge, 551 pp.

Cumming JM (1994) Sexual selection and the evolution of dance fly mating systems (Diptera: Empididae; Empidinae). The Canadian Entomologist 126(3): 907–920. https://doi.org/10.4039/Ent126907-3

Cumming JM, Cooper BE (1993) Techniques for obtaining adult-associated immature stages of predacious Tachydromiine flies (Diptera: Empidoidea), with implications for rearing and biocontrol. Entomological News 104: 93–101.

Edward DA, Chapman T (2011) The evolution and significance of male mate choice. Trends in Ecology and Evolution 26(12): 647–654. https://doi.org/10.1016/j.tree.2011.07.012

Evenhuis NL, Pape T (2019) Systema Dipterorum, Version 2.5. http://www.diptera.dk/

Faircloth BC (2016) Phyluce is a software package for the analysis of conserved genomic loci. Bioinformatics 32(5): 786–788. https://doi.org/10.1093/bioinformatics/btv646

Faircloth BC (2017) Identifying conserved genomic elements and designing universal bait sets to enrich them. Methods in Ecology and Evolution 8(9):1103–1112. https://doi.org/10.1111/2041-210X.12754

Faircloth BC, McCormack JE, Crawford NG, Harvey MG, Brumfield RT, Glenn TC (2012) Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Systematic Biology 61(5): 717–726. https://doi.org/10.1093/sysbio/sys004

Faircloth BC, Branstetter MG, White ND, Brady SG (2014) Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among Hymenoptera. Molecular Ecology Resources 15(3): 489–501. https://doi.org/10.1111/1755-0998.12328

Folmer O, Black MB, Hoeh W, Lutz R, Vrijenhoek RC (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3: 294–299.

Grimaldi D, Engel MS (2005) Evolution of the insects. Cambridge University Press, New York, 772 pp.

Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17(8):754–755. https://doi.org/10.1093/bioinformatics/17.8.754

Ješovnic A, Sosa-Calvo J, Lloyd MW, Branstetter MG, Fernández F, Schultz TR (2017) Phylogenomic species delimitation and host-symbiont coevolution in the fungus-farming ant genus *Sericomyrmex* Mayr (Hymenoptera: Formicidae): ultraconserved elements (UCEs) resolve a recent radiation. Systematic Entomology 42(3): 523–542. https://doi.org/10.1111/syen.12228

Junqueira AC, Lessinger AC, Azeredo-Espin AM (2002) Methods for the recovery of mitochondrial DNA sequences from museum specimens of myiasis-causing flies. Medical and Veterinary Entomology 16(1): 39–45. https://doi.org/10.1046/j.0269-283x.2002.00336.x

Karlsson D, Pape T, Johanson KA, Liljeblad J, Ronquist F (2005) The Swedish Malaise Trap Project, or how many species of Hymenoptera and Diptera are there in Sweden? Entomologisk Tidskrift 126: 43–53.

Katoh S, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30(4): 772–780. https://doi.org/10.1093/molbev/mst010

Lanfear R, Calcott B, Simon YWH, Guindon S (2012) PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution 29(6): 1695–1701. https://doi.org/10.1093/molbev/mss020

LeBas NR, Hockham LR, Ritchie MG (2003) Nonlinear and correlational sexual selection on 'honest' female ornamentation. Proceeding of the Royal. Society B 270(1529): 2159–2165. https://doi.org/10.1098/rspb.2003.2482

Lindahl T (1993) Instability and decay of the primary structure of DNA. Nature 362: 709–715. https://doi.org/10.1038/362709a0

Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harbor Protocols 6. https://doi.org/10.1101/pdb.prot5448

Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Proceedings of the Gateway Computing Environments Workshop (GCE) 1–8. https://doi.org/10.1109/GCE.2010.5676129

Mullen GR, Durden LA (2009) Medical and Veterinary Entomology. Second edition. Academic Press, London, 637 pp.

Murray RL, Wheeler J, Gwynne DT, Bussière LF (2018) Sexual selection on multiple female ornaments in dance flies. Proceedings of the Royal Society B 285(1887): 20181525. https://doi.org/10.1098/rspb.2018.1525

Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. Molecular Biology and Evolution 32(1): 268–274. https://doi.org/10.1093/molbev/msu300

Nurk S, Meleshko D, Korobeynikov A, Pevzner PA (2017) metaSPAdes: a new versatile metagenomic assembler. Genome Research 27(5): 824–834. https://doi.org/10.1101/gr.213959.116

Oosterbroek P (2006) The European families of the Diptera. KNNV Publishing, Utrecht, 205 pp. https://doi.org/10.1163/9789004278066

Orford KA, Vaughan IP, Memmott J (2015) The forgotten flies: the importance of non-syrphid Diptera as pollinators. Proceedings of the Royal Society B 282 (1805). https://doi.org/10.1098/rspb.2014.2934

Poole RW, Gentili P (1996) Nomina insecta Nearctica: A check list of the insects of North America. http://www.nearctica.com/nomina/main.htm

Ratnasingham S, Hebert PDN (2007) BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). Molecular Ecology Notes 7(3): 355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x

Rambaut A (2014) FigTree 1.4.2. http://tree.bio.ed.ac.uk/software/figtree/

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA (2018) Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. Systematic Biology 67(5): 901–904. https://doi.org/10.1093/sysbio/syy032

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19(12): 1572–1574. https://doi.org/10.1093/bioinformatics/btg180

Roskov Y, Ower G, Orrell T, Nicolson D, Bailly N, Kirk PM, Bourgoin T, DeWalt RE, Decock W, Nieukerken E, van Zarucchi J, Penev L (2019) Species 2000 & ITIS Catalogue of Life.

Saigusa T (2012) A new Asio-Nearctic subgenus of *Rhamphomyia* (Diptera: Empididae: Empidinae). Entomological society of Canada 144(2): 291–322. https://doi.org/10.4039/tce.2012.28

Skevington JH, Dang PT (2002) Exploring the diversity of flies (Diptera). Biodiversity 3(4): 3–27. https://doi.org/10.1080/14888386.2002.9712613

Smith TS, Harvey MG, Faircloth BC, Glenn TC, Brumfield RT (2014) Target capture and massively parallel sequencing of ultraconserved elements for comparative studies at shallow evolutionary time scales. Systematic Biology 63(1): 83–95. https://doi.org/10.1093/sysbio/syt061

Ssymank A, Kearns CA, Pape T, Thompson FC (2008) Pollinating flies (Diptera): A major contribution to plant diversity and agricultural production. Biodiversity 9(1–2): 86–89. https://doi.org/10.1080/14888386.2008.9712892

Svensson BG, Petersson E (1987) Sex-role reversed courtship behaviour, sexual dimorphism and nuptial gifts in the dance fly *Empis borealis* (L.). Annales Zoologici Fennici 24: 323–334.

Tagliacollo VA, Lanfear R (2018) Estimating improved partitioning schemes for ultraconserved elements. Molecular Biology and Evolution 35(7): 1798–1811. https://doi.org/10.1093/molbev/msy069

Ultraconserved (2017) Ultraconserved. https://www.ultraconserved.org/

Van Dam MH, Lam AW, Sagata K, Gewa B, Laufa R, Balke M, Faircloth CB, Riedel A (2017) Ultraconserved elements (UCEs) resolve the phylogeny of Australasian smurf-weevils. PLOS ONE 13(9): e0205049. https://doi.org/10.1371/journal.pone.0188044

Wahlberg E, Johanson KA (2018) Molecular phylogenetics reveals novel relationships within Empidoidea (Diptera). Systematic Entomology 43(4): 619–636. https://doi.org/10.1111/syen.12297

Watts M, Winkler IS, Daugeron C, de Carvalho CJB, Turner SP, Wiegmann BM (2016) Where do the Neotropical Empidini lineages (Diptera: Empididae: Empidinae) fit in a worldwide context? Molecular Phylogenetics and Evolution 95: 67–78. https://doi.org/10.1016/j.ympev.2015.10.019

Wiegmann BM, Trautwein MD, Winkler IS, Barr NB, Kim JW, Lambkin C, Bertone MA, Cassel BK, Bayless KM, Heimberg AM, Wheeler BM, Peterson KJ, Pape T, Sinclair BJ, Skevington JH, Blagoderov V, Caravas J, Kutty SN, Schmidt-Ott U, Kampmeier GE, Thompson FC, Grimaldi DA, Beckenbach AT, Courtney GW, Friedrich M, Meier R, Yeates DK (2011) Episodic radiations in the fly tree of life. PNAS 108(14): 5690–5695. https://doi.org/10.1073/pnas.1012675108

Yang D, Yao G, Zhang K, Zhang J (2007) World Catalog of Empididae (Insecta: Diptera). Agricultural University Press, China, 599 pp.

Appendix 1

Table A1. DNA concentration before and after library construction and number of reads for each taxon. Old referring to pinned museum specimens sampled before year 1993, new referring to specimens kept in ethanol and sampled after year 1995. Specimens below lines of average, min and max, refers to specimens excluded from the study.

Taxon	Voucher	Old/new sample	DNA conc. before-library (ng/µL)	DNA conc. after-library (ng/µL)	No of reads
Chelipoda sp.	AH7C	New	0.198	4.16	138000
Empis albinervis	AA9C	New	9.04	2.9	161000
E. bicuspidata	AB1C	New	6.08	3.6	118000
E. borealis	AA8C	New	1.6	3.78	134000
E. latro	AF4C	New	1.07	1.11	40000
E. livida	AA3C	New	0.526	4.22	96000
E. lucida	AA4C	New	0.552	2.46	46000
E. nigritarsis	AD1C	New	4.22	25.6	656000
E. nuntia	AB5C	New	1.21	4.86	265000
E. picipes	AE2C	Old	0.714	13.9	521000
E. planetica	AA7C	New	0.61	2.22	38000
E. tessellata	AA6C	New	0.502	4.04	81000
E. trigramma	AB2C	New	1.12	3.24	109000
E. univittata	AB3C	New	0.892	3.82	150000
Empis sp. 1	AD5C	New	0.894	43.6	1000000
Empis sp. 2	AD7C	New	1.09	35.0	1000000
Empis sp. 3	AF2C	New	0.978	4.7	443000
Empis sp. 5	AG9C	New	0.562	2.42	41000
Empis sp. 6	AH2C	New	1.12	6.14	657000
Hilara cornicula	AG7C	New	0.516	5.1	347000
H. flavipes	AG8C	New	0.486	5.82	294000
Phyllodromia melanocephala	AH5C	New	0.602	7.54	441000
Rhamphomyia curvula	AC9C	New	4.02	18.5	647000
R. dudai	AB7C	New	11.0	1.4	57000
R. erytropthalma	AD3C	New	9.42	18.9	1000000
R. fascipennis	AB9C	New	5.02	2.3	66000
R. gibba	AF1C	Old	0.626	1.58	66000
R. hybotina	AD2C	New	1.48	30.0	1000000
R. nigripennis	AC7C	New	1.62	46.8	3000000
R. nigrita	AD6C	New	0.944	46.4	2000000
R. plumipes	AE6C	Old	1.09	6.02	197000
R. sulcata	AC8C	New	0.868	41.0	1000000
R. trilineata	AB6C	New	2.58	2.72	142000
R. umbripennis	AC1C	New	0.768	1.29	54000
Rhamphomyia sp. 1	AD4C	New	1.52	30.8	2000000
Rhamphomyia sp. 2	AD8C	New	6.62	39.4	2000000
Rhamphomyia sp. 3	AF7C	New	2.7	7.96	721000
Rhamphomyia sp. 4	AF8C	New	1.08	6.58	507000
Average			**2.26**	**12.94**	**558763.2**
Min			**0.198**	**1.11**	**38000**
Max			**11.0**	**46.8**	**3000000**
E. grisea	AE3C	Old	0.688	3.62	110000
E. hyalipennis	AE5C	Old	0.642	1.11	22000
E. nigricans	AD9C	Old	0.518	2.72	176000
E. nitida	AE1C	Old	0.506	0.946	80000
Empis sp.	AF3C	New	0.522	0.612	27000
R. anomalina	AC5C	New	0.362	0.818	19000
R. crassirostris	AC2C	New	0.534	0.584	19000
R. nitidula	AE8C	Old	0.71	0.332	24000
R. pilifer	AB8C	New	1.78	0.138	24
R. spinipes	AE7C	Old	0.672	0.194	28000

Appendix 2

Table A2. Table of partitions and best substitution models generated by Partition Finder v2.1.1.

Subset	Best model	Number of sites	Partitions
1	GTR+I+G	750	uce_1872_core, uce_1165_core, uce_1022_right, uce_1022_left
2	GTR+I+G	176	uce_344_left, uce_1022_core
3	GTR+I+G	376	uce_1165_left, uce_212_right
4	GTR+I+G	270	uce_1165_right
5	GTR+G	415	uce_589_right, uce_1361_left, uce_1361_core, uce_3370_right
6	GTR+I+G	462	uce_2884_left, uce_1361_right
7	GTR+I+G	177	uce_1872_left, uce_2884_right
8	GTR+G	216	uce_1872_right, uce_830_core
9	GTR+I+G	794	uce_830_left, uce_212_left, uce_212_core, uce_589_core, uce_2156_left
10	GTR+G	50	uce_2156_core
11	GTR+G	205	uce_3999_right, uce_2156_right
12	GTR+I+G	99	uce_2884_core
13	GTR+I+G	50	uce_3078_left
14	GTR+I+G	705	uce_3078_core, uce_344_right
15	GTR	142	uce_3078_right
16	GTR+I+G	265	uce_3370_left, uce_3370_core
17	GTR	80	uce_344_core
18	GTR+G	157	uce_3999_left
19	GTR+I+G	652	uce_3999_core
20	GTR+I+G	600	uce_830_right, uce_589_left
21	GTR+I+G	753	uce_716_left, uce_716_right, uce_715_left, uce_715_core, uce_716_core, uce_715_right